LA-UR-79-2895

CONF-791016-4

TITLE: SENSITIVITY ANALYSIS AND A NATIONAL ENERGY MODEL EXAMPLE

AUTHOR(S): Michael D. McKay

MASTER

SUBMITTED TO: 1979 DOE Statistical Symposium, Gatlinburg, TN, October 24-26, 1979.

DISCLAIMER

University of California

LOS ALAMOS SCIENTIFIC LABORATORY
Post Office Box 1663 Los Alamos, New Mexico 87545
An Affirmative Action/Equal Opportunity Employer

Form No. 836 R3
St. No. 2629
12/78

UNITED STATES
DEPARTMENT OF ENERGY
CONTRACT W-7405-ENG. 36

SENSITIVITY ANALYSIS AND A NATIONAL ENERGY MODEL EXAMPLE

Michael D. McKay
Statistics Group
Los Alamos Scientific Laboratory
Los Alamos, New Mexico
October, 1979

ABSTRACT

Sensitivity analysis, a study of changes in a model output produced by varying model inputs, is much more than estimating partial derivatives. As a part of model evaluation, it is an exploratory process directed towards finding out how and why a model responds to different values of inputs. When viewed as a data analysis problem, the intent of sensitivity analysis is to make an inference about a model based on a sample of observations generated from the space of input values. The validity of the inferences is tied closely to the laws, or assumptions, relating the observations (data) and the model.

INTRODUCTION

This presentation is intended to serve as an introduction to a discussion of sensitivity analysis of complex models. By sensitivity analysis, I mean, very generally, a study of changes in a model output produced by varying model inputs. When viewed as a study of variation, sensitivity analysis is much more than estimating what is called a "sensitivity coefficient," the rate of change of an output with respect to an input. Hence, sensitivity analysis may not be the right term to use to refer to the study. Indeed, statisticians, whose primary function is to study and explain variation, are rarely called sensitivity analysts. Nevertheless, I will use the term sensitivity analysis in the broad sense.

The definition of sensitivity analysis - a study of changes in a model output produced by varying model inputs - does not immediately lead to a definition of sensitivity or to definitions of measures of sensitivity. Rather than trying to develop a sensitivity analysis methodology through the definition/theorem approach, I think it is more useful, at least now, to specify objectives of the study - things one would like to find out about the model -

and then look for analysis procedures (sensitivity measures) that will satisfy the objectives.

The rest of the talk will cover some of the thoughts that have evolved during work for the Nuclear Regulatory Commission and the Department of Energy in the area of model (computer code) evaluation. We will look at what is meant by a model, sensitivity analysis as a part of model evaluation, and a data analysis approach to sensitivity analysis. Finally, some results of an analysis of the National Energy Model COAL2 will be presented as an illustration of what can be done in practice.

PART I - THOUGHTS ON SENSITIVITY ANALYSIS

A MODEL

In this discussion a model is a computer program (code) which transforms a vector X of inputs into an output Y(t) which is a parametric function of time. Symbolically, we write $Y(t) = h(t;X)$ where $h(\cdot)$ represents the calculations done by the code. The inputs can be model parameters or problem description variables such as initial conditions. The space of input values for which the model is intended to perform correctly is denoted by S, and t is on some closed interval, say, $[0,T]$.

When models are used to predict what might happen in a real system under some set of conditions - for example, future energy demands or the response of a nuclear reactor to a valve failure - the value of the input vector X which will give the "correct" or best prediction is usually not known. We will describe this uncertainty in the input values by a probability function f(x) on the input space S. Then, a more complete symbolic description of a model is

$$Y(t) = h(t;X)$$

$$X \sim f(x) \text{ for } x \in S$$

$$t \in [0, T].$$

We will assume that $f(\cdot)$ is known, although in practice we may have only incomplete information about it. Often, S is taken to be a Cartesian product of intervals ("error bands" on the inputs) and $f(\cdot)$ is a uniform probability function.

In the following discussions, the parameter t will occasionally be suppressed in the notation for simplicity.

MODEL EVALUATION AND SENSITIVITY ANALYSIS

Model evaluation is concerned with many facets of the $Y(t) = h(t; X)$ relationship, including the logical structure and internal consistency of the computer code, the plausibility of the output for various values of the inputs, and the agreement between the output and experimental or historical data. Since $Y(t)$ is a function of time, model evaluation can be performed for both individual time points and also time intervals.

Sensitivity analysis is the part of model evaluation which studies the changes in a model output produced by varying the inputs. As such, it can cover a very broad area with respect to objectives and techniques. The objectives may include determining the rate of change of the output with respect to each input, ranking the inputs with respect to their importance, or specifying the proportion of output variability attributable to each input.

More simply, one might say that the purpose of sensitivity analysis is to see how a model responds to different values of the inputs. Consequently, one would study a model with the aim of discovering things about it. The key is discovery - of predictable results and unexpected ones, of important inputs and unimportant ones, of a wide range of variability in the output or a not so wide a one.

If we accept the discovery notion, we should view sensitivity analysis as an exploratory investigation until we know what things to look for and how to measure them.

This approach to sensitivity analysis is one of those a statistician can take when studying experimental data. Hence, by looking at sensitivity analysis as a data analysis problem, a statistician might be able to see how to modify and apply statistical analysis methods to the area of sensitivity.

A DATA ANALYSIS APPROACH TO SENSITIVITY

Unless a model has a known and manageable analytic form that can be manipulated mathematically, sensitivity studies will consist of choosing, say, N vectors of the input for which values of the output are obtained. The N pairs $(X_1, Y_1), \ldots, (X_N, Y_N)$ constitute data on which analyses are performed.

As an example, suppose that X_o represents the nominal values of the individual inputs and Y_o is the corresponding nominal output value. Let X_i be the vector of inputs with all values at their nominal except for component i, which has been perturbed by a small amount, say d_i. An analysis could consist of the calculation of estimates of the partial derivatives as $s_i = (Y_i - Y_o)/d_i$. One could use the number s_i as a measure of the sensitivity of the output to input number i.

Viewed as a data analysis problem, the intent of sensitivity analysis is to make an inference about the model h(t; X) based on the observations $(X_1, Y_1), \ldots, (X_N, Y_N)$. Two immediate issues are (1) the selection of input values $X_1, \ldots, X_N$, and (2) the analyses to be performed.

On the selection of the input values, they should be chosen in a manner related to both the intended use of the model and also the expected sensitivity analysis procedures. The selection of the set of input values may be made by intuitive choice, systematic designs, sampling plans, or combinations of these. (The choice of method might be influenced by the cost of the model calculations). Clearly, though, the total merits of the selection procedures cannot be determined without reference to the anticipated analysis.

The analysis of data has as its foundation the observed variation among the outputs $Y_1, \ldots, Y_N$, and the procedure used to select the input values. These two things together constitute the basis for inferences about the model.

Before continuing, something should be said about the difference between data obtained from a computer code and experimental/observed data familiar to statisticians. There is no "random error" associated with computer calculations. Given an input vector X^*, we assume that the calculation $h(t;X^*)$ will always produce the same value $Y^*(t)$. This fact must be kept in mind when talking about things like the variance of an estimator or a t-test.

In view of the properties of the data for a sensitivity study, classical inference (hypothesis testing and estimation) must take on a slightly different perspective. Usually, we make inferences about a population using some kind of a sample from it. The inferences are based on assumptions about the population and the sample. To denote these assumptions, I will use the term "law" rather than the term "model" (as in a linear model) to avoid confusion with the model $Y(t) = h(t; X)$. Generally, laws are assumptions about where the observations come from and how they are related. Laws need not be complete in specification, i.e., $N(\mu,\sigma^2)$ with μ and σ^2 unknown, but they must be complete in the sense that they allow for legitimate inference.

Laws usually include an (assumed) explanation for variation in the data, for example experimental error or different values of an independent variable. It is relative to the laws (assumptions) that analysis procedures are judged with respect to things like optimality, smaller variance, etc.

In summary, the data analysis approach can be viewed as a two step process: selecting model input values and performing analyses involving the inputs and associated outputs. The foundation of the analyses is the observed variation among the output values. Laws (assumptions about the relationship among the input values, the output values, and the model) provide the scope within which inferences about the model are made. The validity and effectiveness of the inferences are judged relative to the laws.

DEFINING IMPORTANCE

Most of the objectives in the next section include some reference to important inputs (or sensitivity). Hence. some quantifiable meaning should be attached to the term "importance," if possible. We might say that an input is important if a change in its value causes a substantial change in the value of the output. A measure of importance could be the partial derivative (sensitivity coefficient). This approach to importance is reasonable when the relationship between the output and the inputs is linear, i.e., under a "linear model" law. Alternatively, one might look at the variance of the output under the linear law, and use the square of the partial derivative multiplied by the variance of the input (or the ratio of this quantity to the variance of the output) as a measure of importance. My preference is for this quantity since it combines rate of change of an input with its range of values, and it is independent of the units of the input.

Propagation of error, as the second technique above is called, can be applied under the law that a Taylor series approximation of $h(\cdot)$ in the inputs X is an adequate representation. In the nonlinear case, however, the partial derivatives are generally non-constant functions of the inputs. Hence, where to evaluate the partials and, even, the possibility of using a directional derivative tend to cloud the issue of quantifying importance.

I want to leave the definition of importance as a possible item for discussion in this workshop. The following points might be kept in mind:

(1) There are many ways to measure importance relative to the observed variation in a sample of output variables.

(2) The laws under which measures of importance are valid should be identified.

(3) The affect on the measures of violation of the laws by the model should be known, at least qualitatively.

(4) Means of detecting violations of laws are needed.

OBJECTIVES OF SENSITIVITY ANALYSIS

In this section, I want to just list some objectives one might have for performing a sensitivity analysis. Having stated objectives, one can assemble techniques and formulate strategies to create a sensitivity analysis methodology.

In sensitivity analysis, one might want to identify:

(1) Important inputs,

(2) Important subsets of inputs,

(3) Important pseudo inputs, for example the product of two inputs,

(4) Important segments of the range of values of an input, for example a threshold value for importance,

(5) Inputs that are conditionally important, for example input A is important when input B is greater than 2.

One could also ask which inputs are

(6) Unimportant inputs.

Some unusual results could be observed in the course of a sensitivity analysis. In this case, an objective would be to

(7) Associate unusual results with specific (subsets of the) inputs or their values.

Certainly, many more objectives could be stated, and each model being investigated will present special considerations. However, even a partial list of objectives can be used to put together a set of techniques and to formulate a beginning strategy for sensitivity analysis.

A QUESTION

What does it mean to say that an input is important, or that the output is sensitive to a particular input?

PART II - AN ANALYSIS OF THE COAL2 NATIONAL ENERGY MODEL

In the following sections I will summarize some of the techniques and results from a analysis of the model COAL2. This work was performed in collaboration with Andrew Ford. from the Energy Systems and Economic Analysis Group at the Los Alamos Scientific Laboratory, and is described in References [1] and [2].

THE MODEL

The model used in our study, COAL2, was developed by Roger F. Naill of the Dartmouth College Systems Dynamics Group under a contract, first from the National Science Foundation, and later from the Energy Research and Development Administration. The model is designed to allow investigators to test a variety of energy policies that may affect the nation's ability to reduce its dependence on oil imports during a period in which domestic production of oil and gas is on the decline.

In our study, 72 inputs were selected for investigation. Of these, 5 were assigned discrete probability distributions. The remaining 67 inputs were assigned uniform distributions on intervals. All inputs were treated as having independent distributions.

Thirteen outputs were recorded at 36 successive time points. The outputs Gross Energy Demand (GED) and Average Energy Price (AEP) will be discussed here.

GATHERING THE DATA

The values of the inputs were selected according to Latin hypercube sampling [3] for the continuous variables. For the sample size of 100, the range of each input was divided into 100 equal (probability) length intervals. The intervals were sampled according to uniform distributions to produce 100 distinct values for each of the inputs. The values of each input were assigned at random and without replacement to the 100 runs.

For the discrete inputs, the values were assigned at random to the runs in proportion to the probabilities of the values.

After obtaining the 100 vectors of 72 input values, we made 100 runs of the COAL2 model. An additional run with all inputs at their mean value was made to generate the base case (nominal) output values.

Plots giving descriptive statistics for the outputs are given in Figures 1 and 2. The statistics were computed independently at each of the 36 time points. Figures 3 and 4 show the output plots obtained for the first 5 runs. These plots typify the data for the two outputs.

From this point our analyses proceed along two parallel paths: formal calculations of a measure of sensitivity as a guide for assigning importance to inputs, and informal examination of the data to get a feel for general trends and irregularities. I will talk here about the formal study. An informal study is found in Ref. [2].

FORMAL STUDY OF THE DATA

The general approach we use in formal sensitivity analysis is outlined in Figure 6. An assumption we operate under is that the determination of important inputs from a set of inputs is easier when the number of inputs is small.

If we let the vector of inputs X represent the set of inputs used to generate the output values, we begin by forming "candidate" subsets of X denoted by $X^c(t)$. At each time t, X is like the variable pool in stepwise regression, and $X^c(t)$ is the subset of the pool that has been selected for inclusion in the regression.

We use step-up partial rank correlation to enter inputs, one at a time, into the candidate subset. We continue to include inputs until (1) the magnitude of the largest partial rank correlation is less than a minimum value, r_s, for selection, or (2) the magnitude of the partial rank correlation for the last selected variable is greater than a maximum value, r_f, which measures the sufficiency of the linear fit.

The results from using several values of r_s ranging from 0.5 to 1.0 are examined to see how the candidate subsets $X^c(t)$ change over time. We have found that using the 95% critical values from the distribution of the ordinary correlation coefficient (with appropriate degrees of freedom) from normal theory produces pleasant results.

The sufficient fit criterion is used in the stopping rule because the number of observations (runs) in our studies is not always greater than the number of inputs. We have only seen this criterion active when the number of runs is less than about 1/3 the number of inputs. The value we use is $r_f = 0.98$.

There are many things that could be said about using partial rank correlation to select the candidate subsets, and about the stopping rules. Certainly, the procedures are ad hoc and depend heavily on linearity to be effective. Let me leave this topic open, and just mention that we look at the difference between results from ranked and unranked data as an indication of nonlinearity.

After constructing the candidate subsets at each time point, two phenomena are often observed: inputs are selected at isolated time points, and conversely, inputs are selected at all but a few of successive time points. Because of these phenomena, we make the assumption that inputs will not be important at only isolated time points. This assumption leads to a refinement procedure for the candidate subsets. We smooth, or filter, the subsets by removing inputs that were selected only once within a time interval of width w. Likewise, if inputs were selected at two time points not further apart than w, we include them at all time points in between.

The choice of a value for w is critical, and requires a study of the results to find a good one. Too small a value for w can produce candidate subsets that vary greatly over time. Too large a value can cause the size of the subsets to get very large. The value used will depend greatly on the modeled event and the time step size relative to the dynamics of the model. In the COAL2 study, we used w ranging from 3 years to 9 years.

Summarizing the first stage of the formal sensitivity analysis, we

(1) From the initial set of inputs X, select candidate subsets $X^c(t)$ at each time point using step-up partial rank correlation.

(2) Filter the subsets $X^c(t)$ to include or remove inputs depending on their occurrence or non-occurrence in neighboring subsets.

After establishing the candidate subsets, we use them to calculate partial rank correlation coefficients as measures of sensitivity. Plots of the determinant of the correlation matrix as a function of time are given in Figures 6 and 7. The plots indicate how well the linear (in ranks) law in the inputs fits the data. The value of 0 indicates a perfect linear fit.

The plots reflect changes in the candidate subsets over time, and somethi of the quality of the laws underlying the analysis procedures. The smooth behavior for GED is easily contrasted with the behavior of the determinant for AEP.

Plots of the partial rank correlation coefficients (PRCCs) are given in Figures 8 and 9.

For GED, a total of 5 inputs were selected over the 36 year time horizon. Input 6, long term growth rate, enters at year 2 and continues as a dominant input through year 36. Input 11, table multiplier for fraction of energy demanded as electricity, is in candidate subsets until year 26. Changes from the value 0 on the vertical axis show when inputs enter or leave the subsets.

We interpret the GED results as indicating that Input 6 is the most important input from year 3 until year 36. Input 11 is important early, and diminishes in importance. After the first 20 years, Inputs 6 and 8 are the only single inputs detectable as important.

A comparison of the PRCCs in Figure 8 with the summary statistics in Figure 1 shows that the size of the candidate subsets decreases as the variance of the output increases over time. One can also see that the analysis and the linear law is less effective toward later years, where more variation is observed in the data. All things considered, past experience directs us to classify results for GED as reliable.

The results for AEP show a somewhat different side of possible outcomes in our sensitivity analyses. The behavior of the PRCCs indicates no clear choices for inputs of standout importance. Input 20 dominates at a relatively high level for the first 9 years. After that time, no inputs have a very high partial correlation with AEP. The conclusion from the PRCCs is that no single variable stands out as important in a high degree (value of PRCC).

A comparison of the PRCCs for GED with those of AEP can lead to the conclusion that the PRCC is incapable of detecting singly important inputs in some cases. Another conclusion, however, is that single variables are not important for AEP, but rather combinations of variables, as in interaction in analysis of variance. In this problem, there are 2556 distinct pairs of inputs. I regret to say that I have not pursued this matter further.

In closing, I have presented the basics of a formal analysis procedure we are using. Sometimes it works very well, at other times it doesn't. It does, however, offer indications of what variables to look at first when trying to explain the behavior of a model.

REFERENCES

1. M. D. McKay, F. A. Ford, G. H. Moore, K. H. Witte, "Sensitivity Analysis Techniques: A Case Study," Los Alamos Scientific Laboratory Report LA-7549-MS, November 1978.

2. F. A. Ford, G. H. Moore, M. D. McKay, "Sensitivity Analysis of Large Computer Models - A Case Study of the COAL2 National Energy Model," Los Alamos Scientific Laboratory Report LA-7772-MS, April 1979.

3. M. D. McKay, W. J. Conover, R. J. Beckman, "A Comparison of Three Methods For Selecting Values of Input Variables in the Analysis of Output From a Computer Code," Technometrics, 21, 2, May 1979.

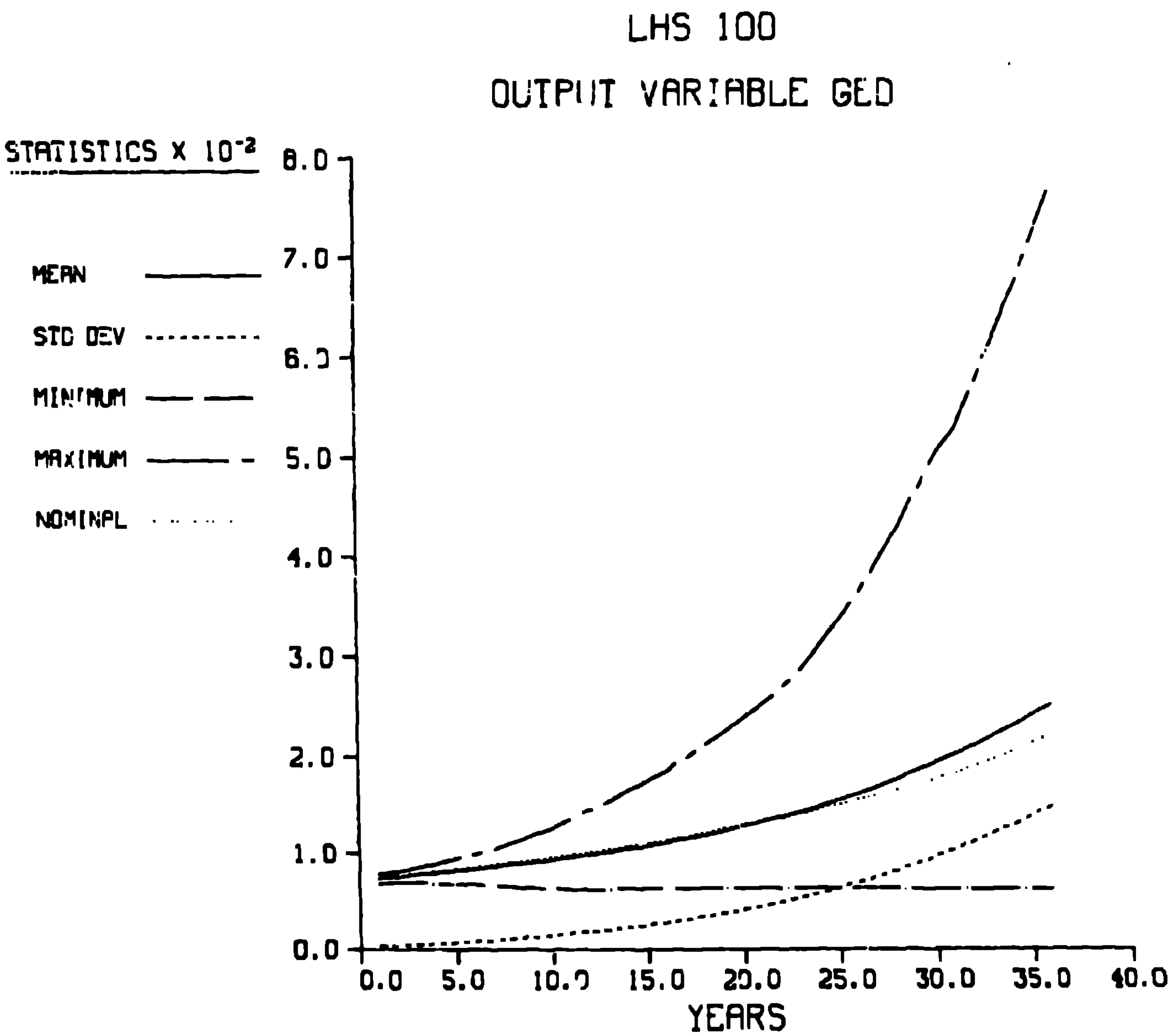

Fig. 1. Descriptive statistics for GED.

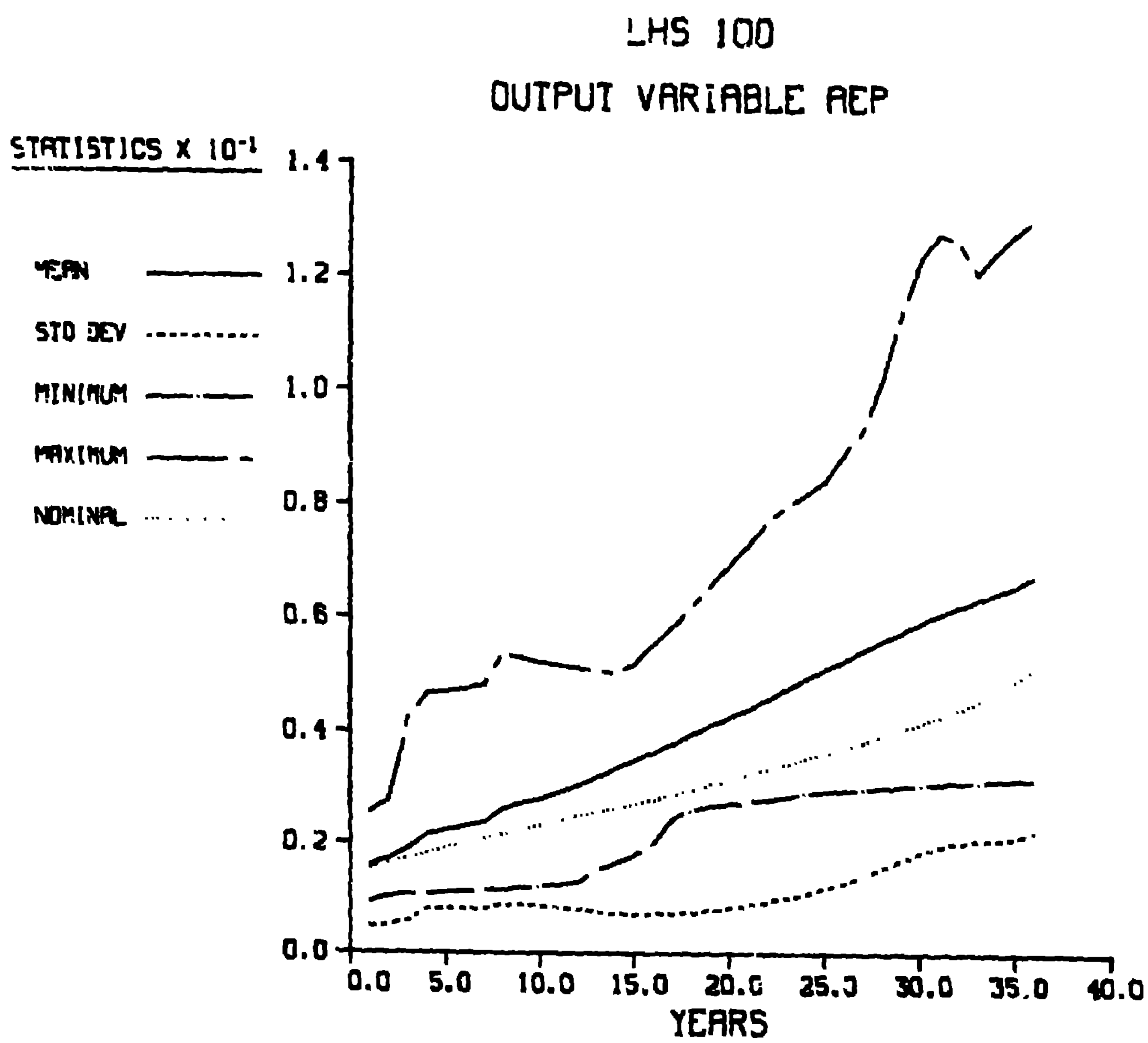

Fig. 2. Descriptive statistics for AEP.

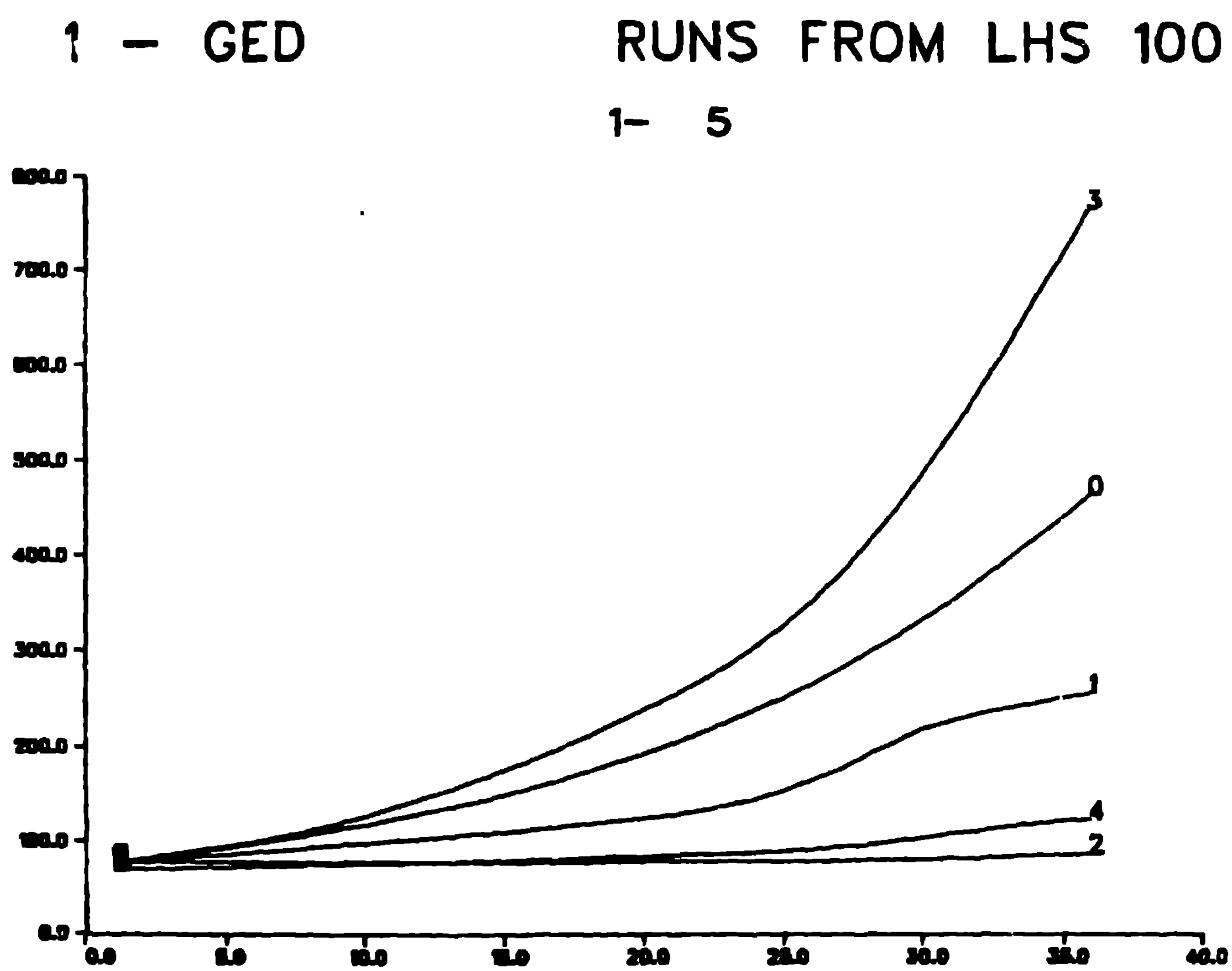

Fig. 3. Data plots for GED.

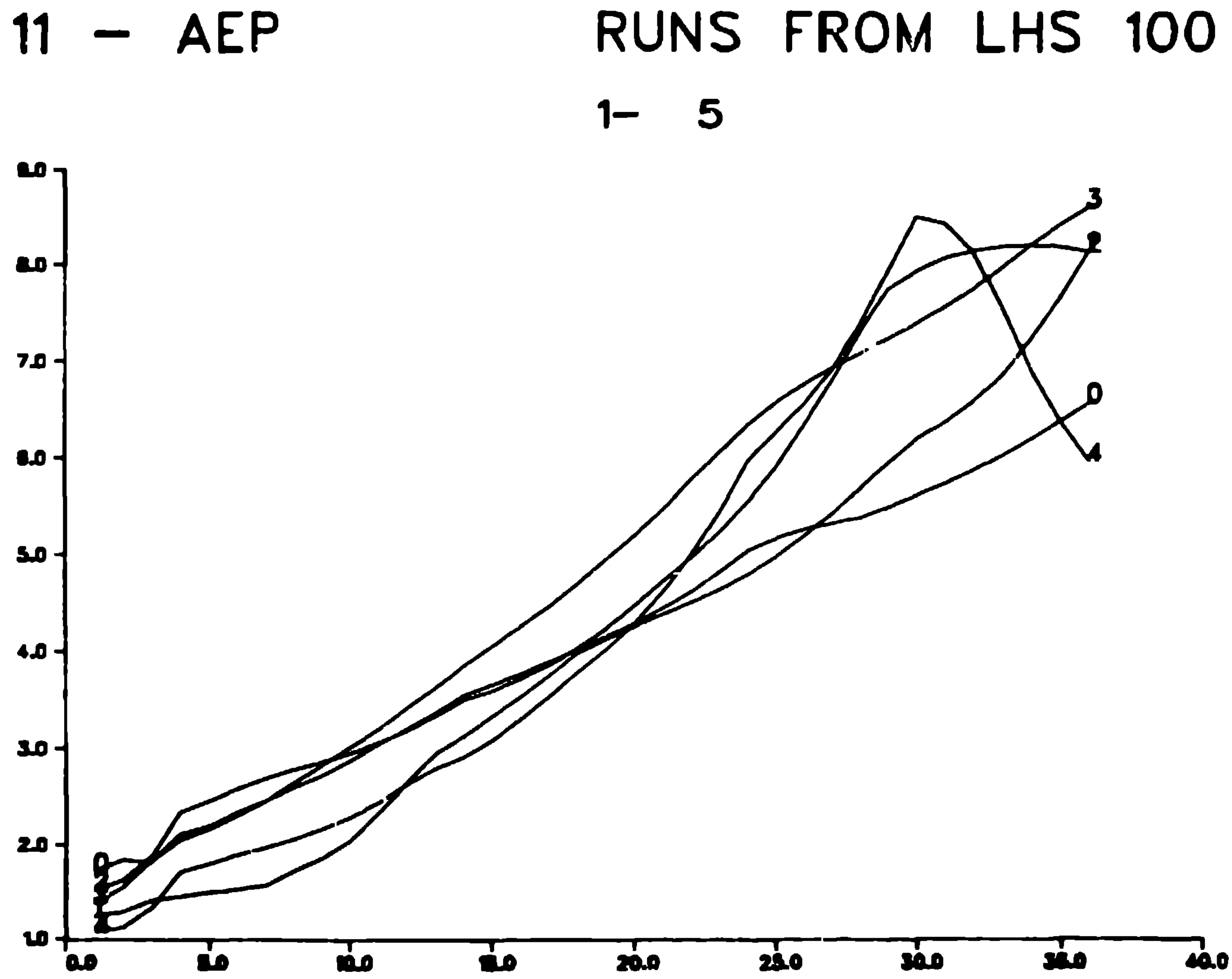

Fig. 4. Data plots for AEP.

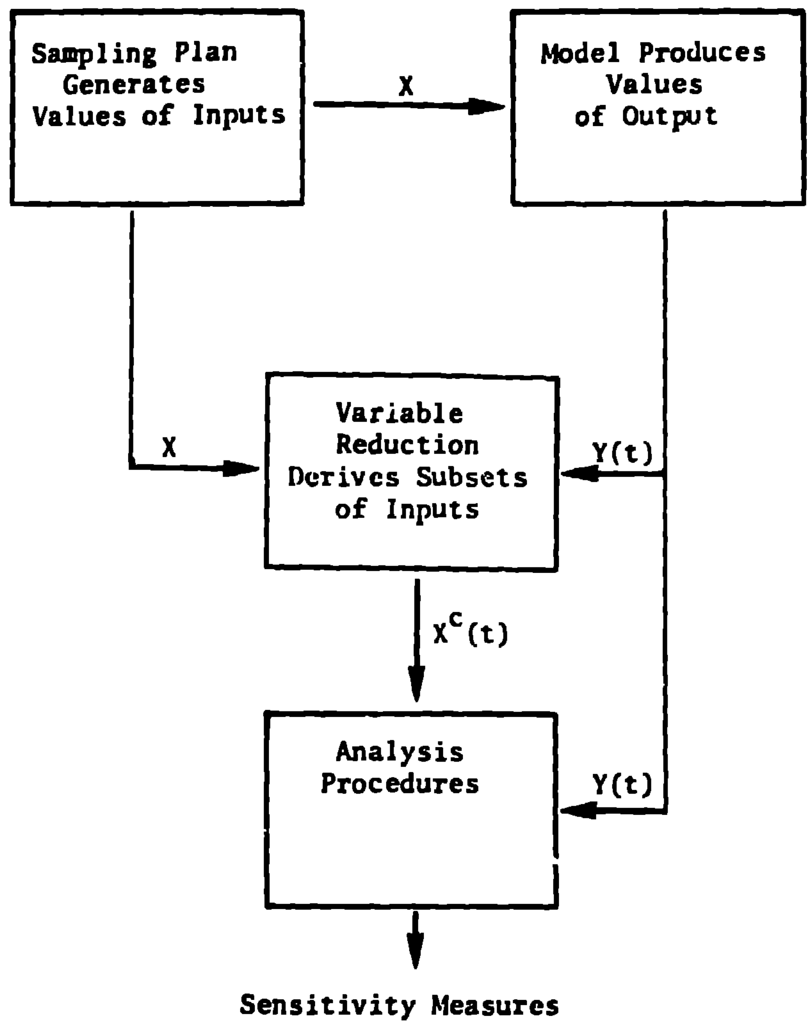

Fig. 5. A formal sensitivity analysis.

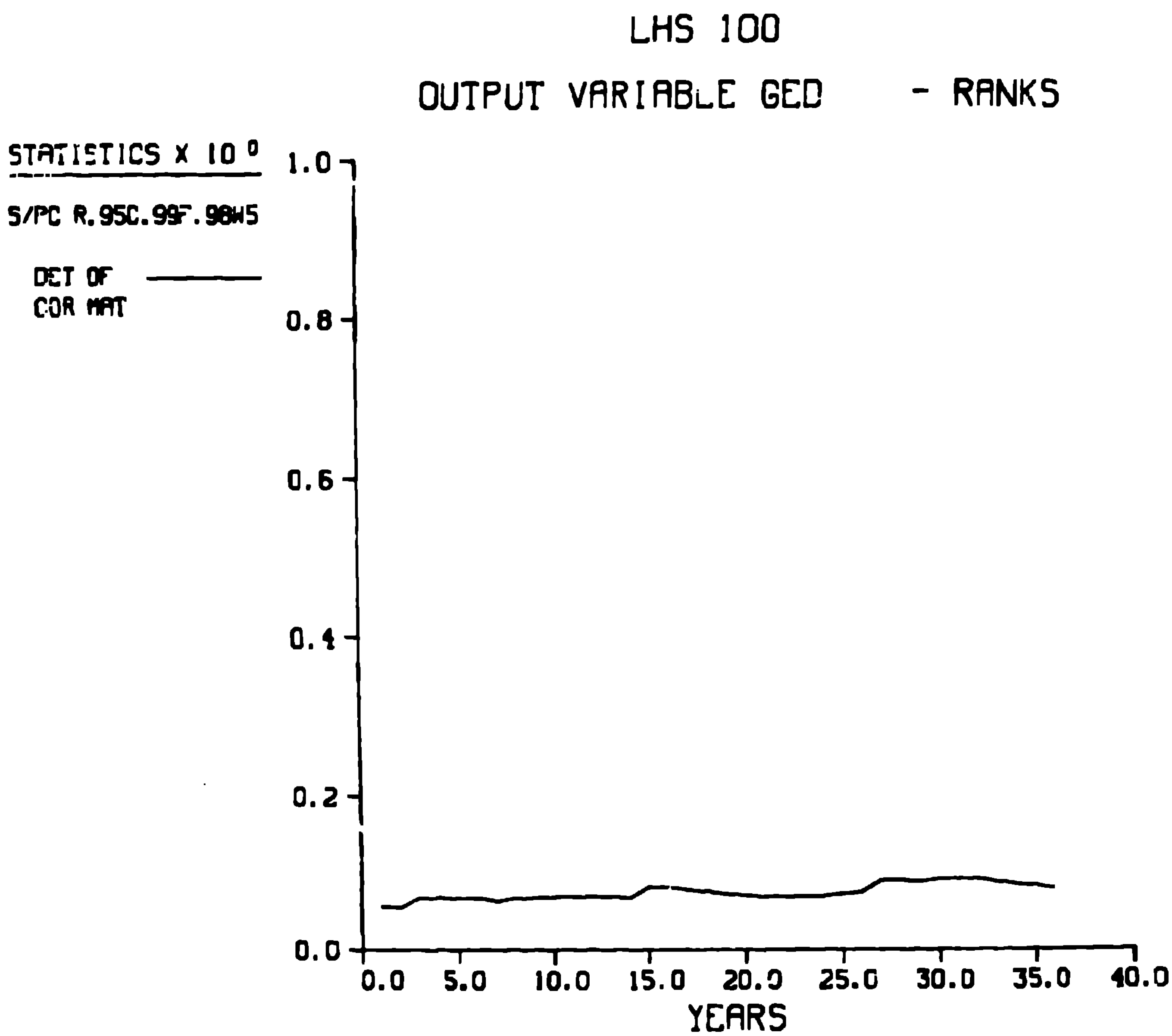

Fig. 6. Determinants of correlation matrix for GED.

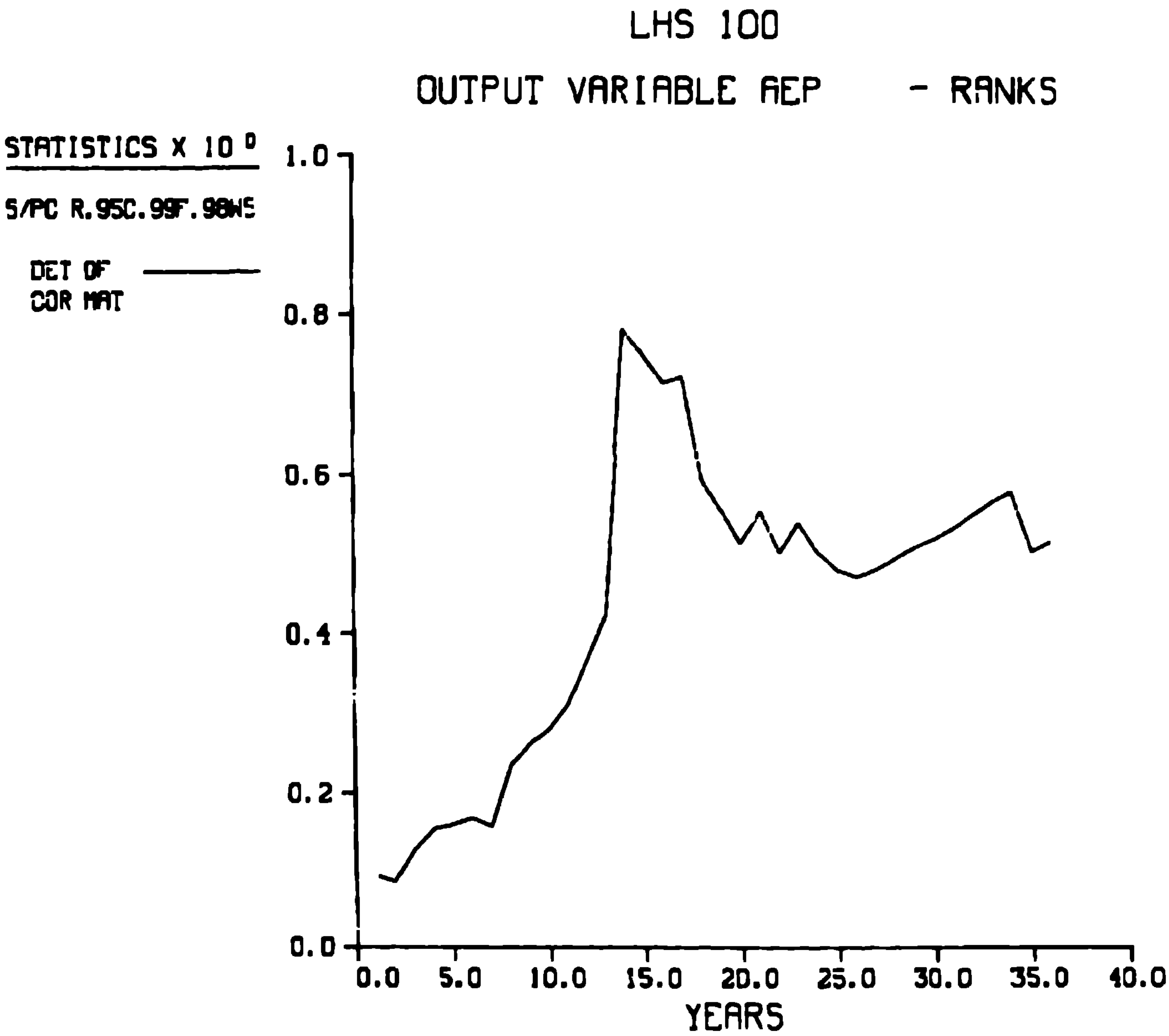

Fig. 7. Determinants of correlation matrix for AEP.

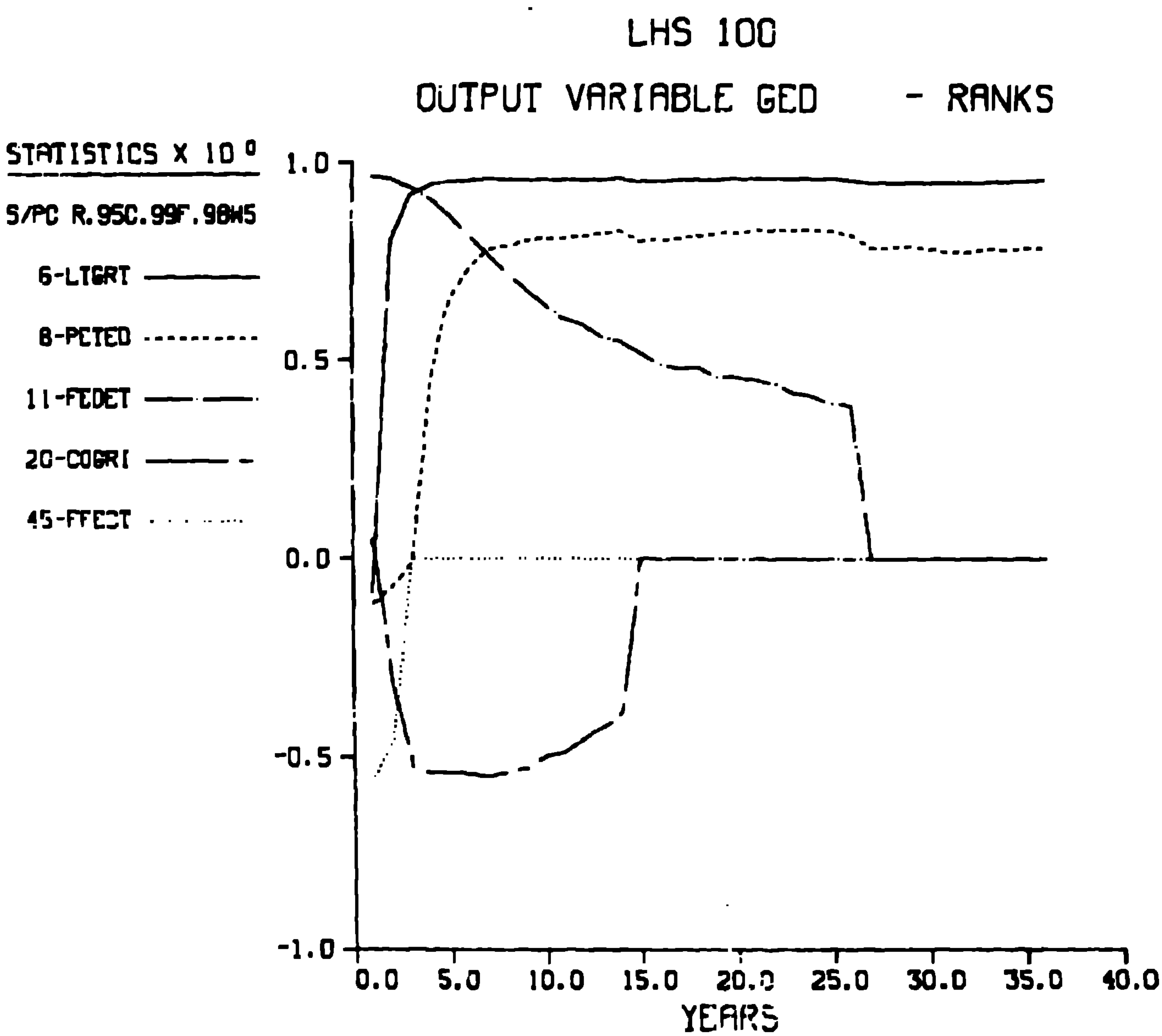

Fig. 8. Partial rank correlation coefficients for GED.

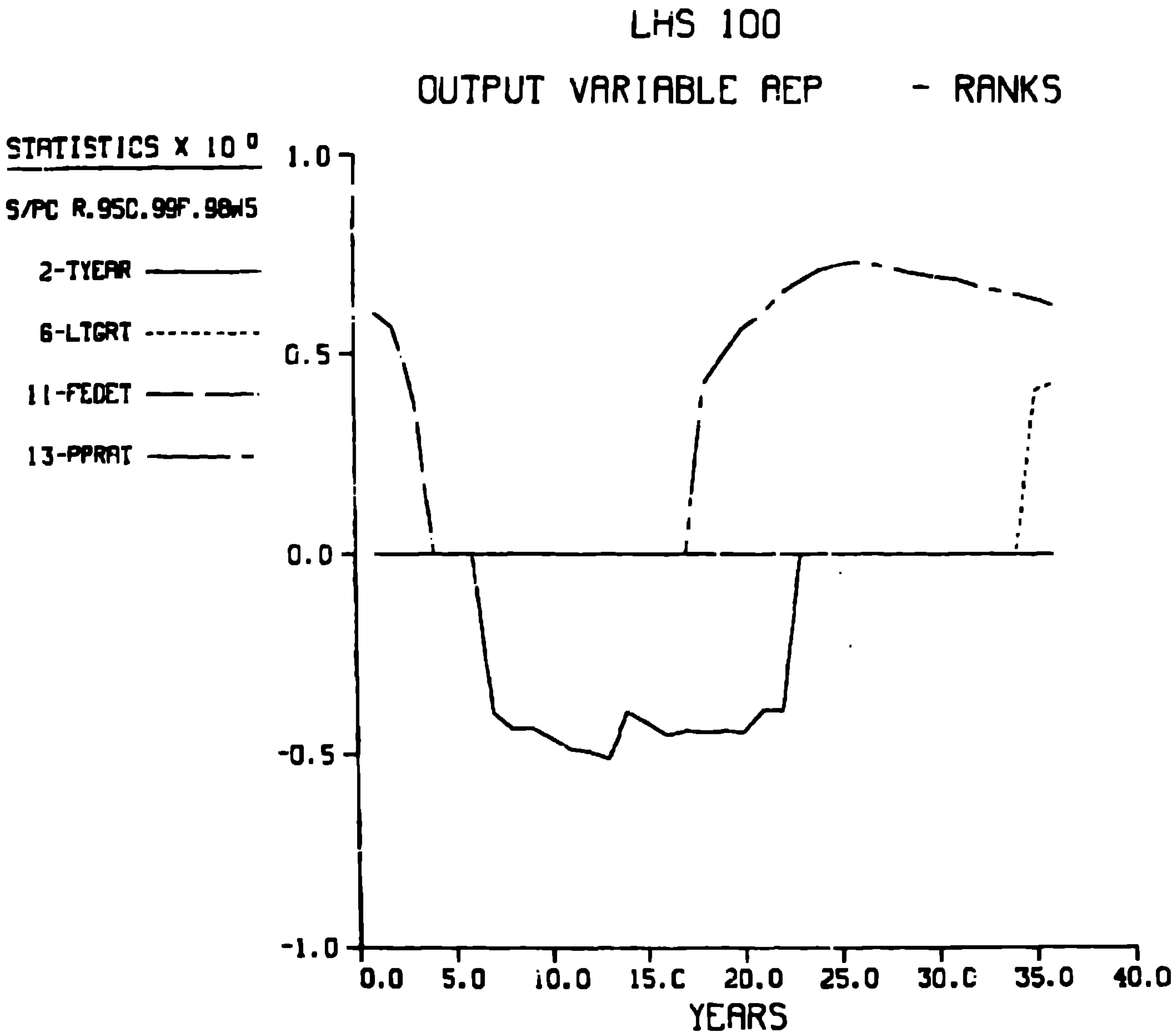

Fig. 9. Partial rank correlation coefficients for AEP.

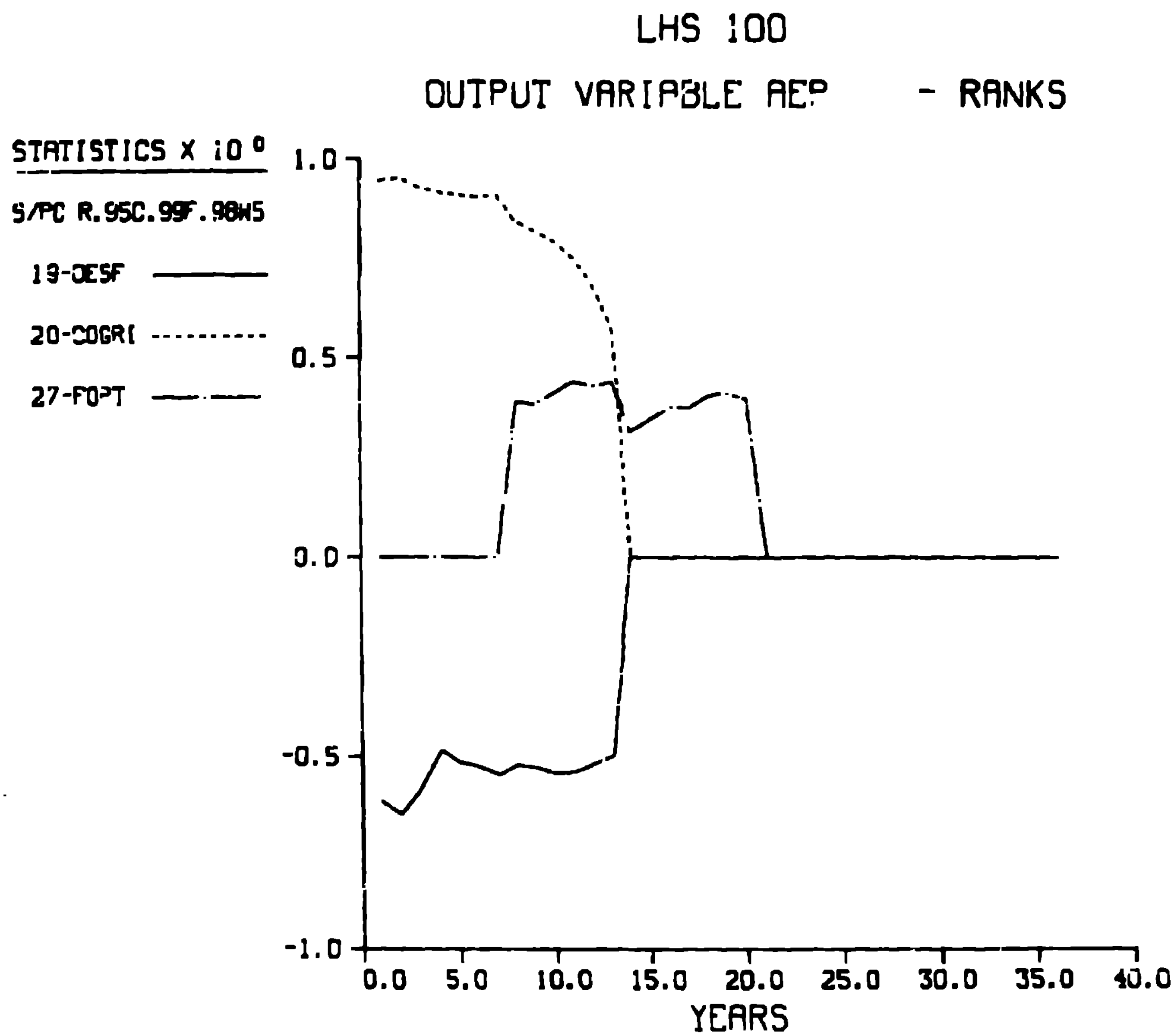

Fig. 9. (Continued). Partial rank correlation coefficients for AEP.